

“It’s Complicated”: Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability

CYNTHIA L. BENNETT, Carnegie Mellon University, USA

COLE GLEASON, Carnegie Mellon University, USA

MORGAN KLAUS SCHEUERMAN, University of Colorado Boulder, USA

JEFFREY P. BIGHAM, Carnegie Mellon University, USA

ANHONG GUO, University of Michigan, USA

ALEXANDRA TO, Northeastern University, USA

Content creators are instructed to write textual descriptions of visual content to make it accessible; yet existing guidelines lack specifics on how to write about people’s appearance, particularly while remaining mindful of consequences of (mis)representation. In this paper, we report on interviews with screen reader users who were also Black, Indigenous, People of Color, Non-binary, and/or Transgender on their current image description practices and preferences, and experiences negotiating theirs and others’ appearances non-visually. We discuss these perspectives, and the ethics of humans and AI describing appearance characteristics that may convey the race, gender, and disabilities of those photographed. In turn, we share considerations for more carefully describing appearance, and contexts in which such information is perceived salient. Finally, we offer tensions and questions for accessibility research to equitably consider politics and ecosystems in which technologies will embed, such as potential risks of human and AI biases amplifying through image descriptions.

CCS Concepts: • Human-centered computing → Empirical studies in accessibility.

Additional Key Words and Phrases: Accessibility, AI, Blind, Image descriptions, Disability, Gender, Race, Visual impairments

ACM Reference Format:

Cynthia L. Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P. Bigham, Anhong Guo, and Alexandra To. 2021. “It’s Complicated”: Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability. In *CHI Conference on Human Factors in Computing Systems (CHI ’21), May 8–13, 2021, Yokohama, Japan*. ACM, New York, NY, USA, 31 pages. <https://doi.org/10.1145/3411764.3445498>

1 INTRODUCTION

Image descriptions play an important role in increasing access to visual information online to people who are blind. As such, web accessibility best practices instruct content creators in the composition of these text-based interpretations, such that they offer a comparable experience to encountering the image visually. Specifically, the Web Content Accessibility Guidelines suggest that quality image descriptions convey the most important information a viewer would garner from the image that is not available from surrounding text [117]. However, best practices lack suggestions for when and how to respectfully describe photographed people’s, called photographees, appearance. This gap is an important area for further consideration by accessibility research, as images of people are key to self-presentation online.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2021 Copyright held by the owner/author(s).

Manuscript submitted to ACM

However, how and when to talk about people's appearances remains contested in popular media and by screen reader users themselves [104]. For example, one blog asked its readers how to describe appearance [23]. In another example, the audio description—narration of imagery in videos or live action—of the film *Black Panther* garnered critique as a white, British audio describer was cast to narrate, and ostensibly choose what was described and what was left out about a film meant to be unapologetically Black [76, 89]. Some popular sources converged that describing race, gender, and disability of people in images is sometimes important and appropriate if the photographee has willingly shared this information [21, 67]. But whether a photographee's appearance is important enough to be described is often decided by describers, who are often not themselves screen reader users. In turn, accessibility scholars have published what screen reader users want from their image descriptions [73, 104]. We expand conversations about human descriptions in image captions by inquiring how appearance is negotiated by blind people, as well as when and how screen reader users believe image descriptions may respectfully communicate the appearance of photographees.

Meanwhile, accessibility researchers are innovating methods of proliferating image descriptions, which remain scant despite the afore-mentioned image description best practices [39]. As such, accessibility research has prioritized applications to increase their presence. For example, image descriptions created using machine learning methods, which we refer to as AI-generated image descriptions, pose a promising avenue to solve image accessibility issues given their potential to scale while reducing the burden on content creators [123]. However, as demonstrated by recent policies limiting AI's usage in public [52], and numerous research contributions including [17, 95], AI-powered analysis of faces, bodies, and associated data has come under scrutiny for promoting harmful biases and surveilling already marginalized people, such as Black and transgender people. Yet, work at the intersection of AI, accessibility, and ethics remains novel and narrow; many popular conversations on AI bias tend to propagate AI as an accessibility solution, rather than a potential risk. For example, some currently-available products, like Seeing AI, share AI-generated appearance information, like the age and gender of photographees, to end users when they upload images to the service [70]. As research continues to innovate in producing image descriptions, we contribute insights on respectfully describing appearance, and possible modes of description generation and implementation (human, AI, and others). Specifically, we offer four contributions:

- First, we offer perspectives by people very under-represented in HCI and accessibility research, and who represent an interesting nexus of this research area. Participants were screen reader users, a target user and assumed beneficiary of image descriptions, including those which may be AI-generated. But participants also identified as being either or both a minoritized race or gender which has been shown to be misidentified and negatively impacted by AI, such as being Black, Indigenous, a person of color, non-binary and/or transgender. When researching the impact of technology, with few exceptions like [29], HCI and accessibility scholars tend to inquire populations that occupy one axis of marginalization [88, 98]. Yet AI bias analysis demonstrates that inquiring intersections of marginalization may point to unique and consequential impacts [17], motivating our focus.
- Second, we report how a subset of screen reader users negotiate appearance nonvisually. These experiences push back on assumptions that vision is a universally-available form of communication through which race, gender, and disability may be presented.
- Third, we inform image description best practices with preferences for when and how race, gender, and disability may be described. These insights serve not as guidelines but as evolving considerations and questions.
- Finally, we share user perspectives on AI and bias by probing their thoughts on the ethics of AI-generated descriptions of appearance.

In what follows, we background disability, race, and gender before reviewing three key areas of literature that underpin our work: self-presentation online through images, image descriptions, and AI bias. We then detail our interviews with 25 screen reader users, all of which identified as at least one of the following: Black, Indigenous, person of color (BIPOC), non-binary, or transgender. We then discuss the findings in the context of AI's application to increase access to visual content. In summary, participants wanted to learn more appearance information about photographees from image descriptions than they currently receive, but they were wary of this information being written into all image descriptions, and they agreed these descriptions should be written with care. As such, they shared contexts where knowing appearance information may be particularly helpful, and they offered language that may be acceptable for describing appearance depending on what a describer knows about the photographee's identity. Finally, participants shared their excitement and concerns about potential benefits and harms of biased AI-generated image descriptions. As such, we discuss cautions in deploying them. In particular, we call for greater accountability and design decisions which take seriously mitigating potential harms of AI for accessibility, so users are not forced to depend on potentially harmful technologies to gain critical access.

2 BACKGROUND

2.1 Definitions

In this paper, we aim to better understand how screen reader users negotiate appearance nonvisually and appearance's role in conveying race, gender, and disability. As such, we briefly background these phenomena to align readers with our understanding, as they remain widely contested. Specifically, we discuss them in the context of power circulation, wielded structurally and individually, which produces uneven but predictable oppression and opportunities depending on how one identifies with and is coded into subcategories via systematic ableism, racism, and transphobia, respectively. Our position draws upon U.S. understandings which aim to foreground rather than disappear social and material consequences of these identity facets, detailed next. In this section, we describe (1) disability in the context of screenreader users; (2) gender; and (3) race.

Screen reader users in this work comprise people who use screen reader software to complete most computing activities. Generally, screen reader users have visual impairments that make completing computing tasks visually uncomfortable or impossible. Screen reader users, or people who are blind or visually impaired, are more generally classified under the umbrella of disability. Disability refers to the margins of the sociomaterial system of normalcy that orders bodies and minds according to normed expectations of how they should exist and behave. Disabled people comprise those who are classified outside such norms and those who have enfolded those experiences into their disability identity. But whether or not someone claims a disability identity, and the language they use to do so varies greatly, making the category important to study in the context of how it may be used in image descriptions. Ableism then, is prejudice that reinforces normed body and mind hierarchies and the erasure of disabled people and their experiences. Ableism does not just impact people who identify or appears as disabled but anyone who does not fit these expectations [30, 31, 61, 66].

Gender refers to synthesis---or lack of synthesis---of one's embodied physical form, their internalized conceptual form, and/or their social presentation of a gendered identity. Some people's genders are different from those assigned at birth, and they may identify as transgender, non-binary, among many others. Cisgender, in contrast, refers to people whose gender is the same as the one they were assigned at birth. Transphobia refers to prejudice against and erasure of transgender people. Given pervasive prevalence of transphobic assumptions that gender is binary and fixed, people

who are not cisgender are disproportionately misrepresented, and such misrepresentation is further used as a tool of harassment and harm. As these highly personal but contested gender labels could be included in image descriptions, we explicitly recruited screen reader users whose gender is different from that assigned to them at birth, such as those who are non-binary and transgender, to learn how gender and misrepresentation impact their experiences and preferences around image descriptions of appearance [36, 96].

Race is a sociomaterial system of categorizing people based on factors including, but not limited to, ancestry, culture, physical features, social and behavioral qualities, and self-identification. Racism comprises attitudes, action, and institutional policy and structure that subordinates people due to their race [60, 108]. Racism can occur at institutional, individual, and internalized levels. Racial coding and frequent misidentification also disproportionately impact BIPOC people as racism perpetuates these (mis)labels (which could be included in image descriptions) to lead to inaccurate and unjust beliefs and behaviors towards them [82, 92].

Together, these brief definitions are not meant to serve as comprehensive, but we provide them to anchor forthcoming discussions. Finally, though we define these terms separately, we note that they are always intersecting and entangling to produce unique experiences and perspectives [25]. That is why, for example, we recruited people who are minoritized by prejudice against their disabilities as well as their races and/or genders. Screen reader users who are BIPOC, non-binary, and transgender may benefit from image descriptions, but they are also among those more often misidentified by AI. As such, they comprise a relevant but understudied nexus of perspective for this research topic.

2.2 Nonvisual Sensemaking of Appearance

Blind people encode both visual and nonvisual understandings of disability, gender, and race while interpreting appearance, identity, and difference [63, 81]. But appearance and identity are often assumed to be primarily visual phenomena. For example, the ableist “colorblind” metaphor for ‘not seeing’ race incorrectly gives rise to an understanding of race that if we were only able to visually ignore or omit race that racism would be abolished [4, 41]. Instead, the perception and visual salience of race are social phenomena that pervade our society, including those with visual impairments who live within it [81]. In a recent example, prejudice was pointed out among blind people themselves, materializing the ways an absence of vision does not preclude narrow and harmful actions that are often assumed to stem from visual perceptions of difference. A presentation by Black blind activists at the National Federation of the Blind’s 2020 convention detailed their numerous racist encounters with other blind people, prejudice rooted in both nonvisual interpretations of their language and behavior, as well as visual cues they picked up from others [49]. That said, separate interviews with blind people revealed great uncertainty about the nonvisual methods they leverage to judge race and gender [33]. Conflation of vision with power to know has led blind people, and others more broadly, to downplay the role of nonvisual sensemaking of race, gender, disability, and other attributes from appearance while overstating the validity of visual perception as definite and confirmatory [33, 63, 81]. For this paper, these lessons indicate that understanding appearance is important to screen reader users, and screen reader users are not immune from impacts of, and from participating in, meaning-making of appearance and associated ableist, gendered, and racist prejudices.

3 LITERATURE REVIEW

Below, we briefly overview key literatures of this work. These reviews of self-presentations online, nonvisual image descriptions and critical perspectives on AI and human analysis are meant to be touchpoints into our grounding rather than comprehensive.

3.1 Self-Presentation Online

Identity and technology in the 21st century are closely intertwined. Both individual and community notions of identity are constructed through how people self-present online [9], online discourse, culture, and practice (e.g., Asian American and Pacific Islander (AAPI) identity work as deliberation [28], Black Twitter [16, 83], queer visibility and self-presentation [18], and re-conceptualizing socialization for autistic children in Minecraft [91]). Identity simultaneously is personal and contextual as well as linked to systems of power, privilege, and oppression. Historically, a dominant narrative insisted that identity was unimportant, invisible, and irrelevant in digital and technological contexts (e.g., “on the Internet, no one knows you’re a dog” [32]). Today, instead we understand that identity pervades our data in both visible and invisible forms.

Specifically, posted photos, such as profile pictures, have come to hyper-symbolize self-presentation. Researchers, for example, have analyzed profile pictures to judge personality traits, and candidate photos may impact hiring decisions [7, 56, 100]. Additionally, social media users both intentionally present and hide their race, gender, and disabilities through their photos [19, 62, 71, 85, 126]. Finally, posted images themselves are not only important in negotiating race, gender, and disability, but the comment threads they facilitate also become sites of identity management [90]. Though self-presentation is not entirely visual, as information included on a social media profile certainly contributes, photos are highly influential in communicating appearance and identity.

3.2 Nonvisual Accessibility of Images Online

For screen reader users to access images online, the image must be accompanied by an image description, often called alternative text that is read aloud by screen reader software when the image is encountered [116, 117]. Ensuring that images on the web contain descriptions, especially high-quality ones, has been a perennial problem in accessibility research since the dawn of the web [47, 48]. However, access to images is important to blind people, who aim to interact with them as their sighted counterparts, keeping this research focus a priority [12, 40, 68, 73, 115, 122–124]. Image accessibility has undergone large improvements over the past decades; described images on large commercial and news websites increased from approximately 40% in 2006 [14] to 72% in 2018 [43]. However, this increase has been countered with a rapid increase in user-generated content on social media sites, content which is often non-textual and inaccessible to screen readers [39, 68, 74]; only 0.1% of images posted on Twitter appeared to have image descriptions in 2019, for example.

Researchers have studied several solutions to increase the number of described images from crowdsourcing [93] to generating them automatically [40, 123]. These solutions have brought some benefit; 81% of crowdsourced image descriptions were deemed to be relatively high quality in one study, and users of Facebook’s automatic alt text appreciated the key word descriptions to help them determine whether to take time to obtain a full image description from another source. However, current automatically-generated image descriptions lack the detail and accuracy sought by screen reader users. Progress to encourage content creators to write descriptions are hampered by design choices which hide [39] or do not support image descriptions altogether (e.g., Snapchat, Pinterest, Reddit, and TikTok). Additionally, there is some evidence that screen reader users may over trust AI-generated image descriptions, raising the need for more public education about AI’s limitations [68, 124].

As research continues exploring how best to scale image descriptions across the internet, so too have scholars inquired what screen reader users want to learn from image descriptions in different contexts. The recommendations coming from this literature echo best practices that image descriptions should prioritize the most important information.

But what has found to be important varies from objects to people to actions, scenes, colors, image quality, and image type (e.g., close-up) [73, 84, 104, 105, 115, 122, 124]. Regardless of the context, however, photographees are usually perceived as interesting by screen reader users. Researchers have garnered that the number of people in an image, facial expressions, and whether a photo is a selfie are important details [124]. Stangl *et al.* [104] are among the few scholars who have begun to explore what screen reader users want to learn about the physical features and identities of photographees. In their study, participant preferences diverged on whether and how aspects of appearance like race should be represented. Viewpoints included that race was important to describe if the image was related to a social justice topic; others wanted access to a photographed person's race any time, believing it as part of basic access. Finally, others questioned the possibility for an image describer to judge race accurately. In this paper, we deepen this preliminary inquiry into screen reader users' preferences for how appearance should be described.

3.3 Critical Perspectives on AI and Human Analysis

Identity is continuously embedded into machine learning and artificial intelligence work, from natural language processing (NLP) to computer vision, both critical domains in automated image captioning. Often, identity characteristics like race and gender show up in two ways: explicitly and implicitly. For example, NLP might rely on explicitly gendered features, like gendered pronouns or terms (e.g., actor versus actress). Computer vision might similarly rely on explicitly labeled identity characteristics. For example, a variety of facial analysis tools classify race and gender based on visible difference, thus encoding specific notions of identity into algorithms [95]. Many have critiqued the very labels being used to encode such identities into algorithms as socio-historical, situated and thus incomplete, inaccurate, and subjective [51, 97].

On the flip side, identity information may be inferred, by proxy. Often, machine learning researchers omit identity data in an attempt to remain universally objective or to avoid bias. Yet even when demographic categories are excluded from training data in attempts to ignore social difference, algorithms have been known to replicate identity-based biases [10]. A well-known example, among many others, includes Amazon's scrapped resume parser, which relied on natural language processing to funnel applicants to recruiters but wound up privileging male candidates through gendered language proxies [114]. For marginalized groups, the misclassification of identity is a form of violence and reflective of greater societal discrimination [11, 17, 50, 53, 54, 64, 99]. The power to identify sensitive characteristics like race and gender have material, real world consequences such as the surveillance of Uighur people [75], wrongful arrests [2], and nonconsensual use of images in datasets [102]. In response, critics have called for required consent before facial analysis is performed on them or their photos [38]. But some Black and trans people have already voiced their disinterest in being subject to facial analysis [94, 95], for example. Recently, commercial purveyors of AI classification systems have adjusted their product offerings to not expose gender labels [37] or by redacting facial recognition or analysis software [46, 58, 109].

Accessibility is unique within AI fairness conversations as AI is treated as a promising tool to automate access and increase accessibility at scale. Its potential applications for people who are blind or low vision alone run the gamut from computer vision applied to recognize known people in one's vicinity [42, 106, 125] to increase blind navigators' feelings of awareness and safety [15]. Additionally, research has shown that passersby may be more sympathetic to the use of surveillance technologies, like those which may feed AI-augmented analyses of surroundings if their purpose is known to assist someone with disabilities [86].

But with these potential benefits come similar concerns about bias. As such, accessibility and AI researchers have convened positions, workshops, and recommendations for remediating and reforming AI towards more fair solutions

including [59, 113]. Researchers have raised awareness of unique impacts of AI on people with disabilities, who are often considered outliers, and therefore left out of data altogether [44, 77, 111]. Through this education, we learn of disability's complexity. First, not everyone discloses their disabilities. Second, as a protected class, collecting health and disability data remains difficult yet such data could help in the accessible customization of technologies, and disabilities are extremely diverse, making any collective recommendations difficult [72, 112]. In response to these and wider critiques of biased AI, scholars have published paths forward. For example, recommending more representative and transparent data, intentionally designing for outliers, and methods which minimize data sharing [3, 34, 35, 44]. Still, others warn that the allure of accessibility as an altruistic application of AI, even if fair, may excuse its perpetuation of the afore-mentioned harmful surveillance and misclassification [11, 118–120].

AI fairness remains a splintered specialization where concerns are often considered singularly, such as fairness that may reduce only one type of bias at a time. Further, these perspectives have not delved into the specifics of appearance classifications in image descriptions. This gap may lead to image description solutions which neglect first person perspectives on sensitive aspects of appearance that, if described inappropriately or incorrectly, could cause harm to people already shown to be disadvantaged by AI and greater society. As information about photographees is valuable when screen reader users browse images, understanding how appearance is negotiated by multiply marginalized blind people and their preferences and concerns around AI labeling such information is imperative.

4 METHODS

The 25 participants recruited for this study were comprised of adults who used screen readers for most computing tasks, including the social media browsing they did at least several times per week. Additionally, they identified with at least one of the following: Black, Indigenous, person of color, or not cisgender.

Given privacy concerns, the first author recruited participants through her connections in the blind community, reaching out specifically to leaders of activist groups comprising members who are blind BIPOC and LGBTQ+. We recognize this method of recruitment stemmed limitations in that our participants tended to be more open to talking about race, gender, and disability, had pride in their identities, and they possibly had more awareness of AI and fairness and accessibility activism than a random sample. We chose this method so we could intentionally recruit people with a variety of lived experiences and so we could have direct conversations with potential participants to discuss and address any concerns. In particular, we note this was important for building trust among and subsequently recruiting LGBTQ+ individuals who were in various stages of disclosing their gender to different audiences. Participants were aware they would be asked to discuss race, gender, and disability before the interview, and that questions may be sensitive. We clarified they could decline answering any questions or end the interview at any time.

Interviews were between 60-90 minutes. With permission, interviews were audio recorded and all participants were compensated for their time. Questions concerned five categories:

- Each interview's outset concerned participants' identity and their experiences with misrepresentation. Specifically, we asked them to describe their race, gender, and disabilities and to share how the language they used to describe these identities changes, if at all. We then inquired whether they have been misrepresented—if someone else has made incorrect assumptions about their identity.
- The second part of the interview concerned participants' social media use and their image browsing and posting behaviors. We recognize that image descriptions are needed across the internet and this focus limits our study.

However, we dialed in to this domain as social media was the locale of most participants' intentional photo browsing and image description composition.

- Third, we asked participants about their behaviors and preferences around disclosing their own appearance, and learning others' appearances. We began by asking them to share situations in which they disclose their appearance in image descriptions (if they did not share this information earlier). The conversation continued as researchers inquired into situations in which they are interested to learn the appearance of others, their strategies for obtaining this information nonvisually, and remaining challenges.
- Fourth, we asked participants to share which language they prefer be used when their appearance is described both by image describers who know them and who are strangers.
- The interview concluded with a conversation on AI-generated image descriptions. We began by learning of any experiences participants had encountering them, including any experiences with misrepresentation. We then scaffolded a conversation to understand how their preferences around AI describing aspects of appearance in image descriptions changed, if at all. We began by asking for positive and negative reactions to AI describing photographed people's appearance that may indicate their race, gender or disabilities. If participants did not talk about AI bias freely, we then asked if they were familiar with known instances of AI bias such as that shown in the Gender Shades project where darker-skinned women were more often algorithmically misclassified than lighter-skinned men [17]. Once the topic of bias entered the conversation, we asked participants to share their positive and negative thoughts by weighing the potential benefits of increased access to image descriptions with potential risks of AI generating them.

The first author transcribed the interview recordings and thematically analyzed quotes according to prominent and unique topics discussed. As transcripts accumulated, the research team reviewed, iterated, and arbitrated thematic memos until reaching consensus on those most salient for this work. Specifically, analysis began deductively. Quotes were divided according to the five interview topics shared earlier. Transcripts were then analyzed for inductive themes [20, 107]. For example, we intentionally looked for quotes concerning experiences being misrepresented, but inductively, we determined that these experiences usually concerned misrepresentation of their race and gender, instances perpetrated both outside and inside of the blind community, and some instances were particularly impactful whereas others did not bother participants.

The ethos of interviews shifted from enthusiastic to difficult. Participants were generally excited to think with us about their identities and appearances in the context of image descriptions and AI, as they found the topic under-discussed, summarized by Aqua, '*I think [blind] people feel weird about talking about this; they're like, you shouldn't just be [nonvisually] assuming people's race. That goes on [visually] all the fucking time.*' But they also vulnerably shared experiences of marginalization while circling with moments of reckoning that in describing how they nonvisually learn and assume others' appearance, they relied upon problematic stereotypes that build the hierarchies that suppress their own identities. As researchers, we analyzed transcripts with reflexivity that our presence and responses influenced each interview [8]. For example, the first author could use her identity as a blind person to build rapport with participants, but as a white cisgender woman, she needed to remain an active listener on topics of race and gender. Further, we grappled with how to report participant experiences of and views on prejudice with the aim of not compounding their harm in sharing them secondhand. As such, the findings contain some paraphrases over the detail of direct quotes in attempt to negotiate these topics with care. Altogether, we present findings that, we hope, can move important conversations on respectful image descriptions.

5 FINDINGS

5.1 Demographics and General Social Media Usage

We report aggregated demographics to protect participant privacy. Seventeen participants had visual memories identifying physical features of people, and 14 were totally blind at the time of the interview. Five participants look at photos on social media though all clarified their preference is to engage with a photo by reading an image description with their screen reader. All but one participant ranged in age from 20 to 42; the 25th participant was 60 years old ($M = 32.63$, $SD = 7.98$). Seventeen participants were BIPOC and four of those have white passing privilege. Eleven participants were not cisgender. Three participants were both BIPOC and not cisgender. Specific races and genders of participants are reported in tables 1 and 2. Finally, 10 participants shared disabilities other than blindness, most disclosing more than one.

Social media use and photo engagement largely aligned with prior work [39, 40, 73, 115, 123]. For scoping purposes, we report on social media use that was not related to online dating, as that topic ended up not being prominent in our interviews. Briefly, participants found online dating still largely did not support basic accessibility. Almost all (23) participants used Facebook several times per week; other frequently-browsed accounts were hosted on platforms including Twitter (14 participants) and Instagram (6 participants). All 25 participants cited the presence of a human-written image description was the most likely predictor that they would engage with an image on social media. Though several participants mentioned this behavior was rare, they shared strategies they employed if they are committed to learning more about an undescribed image. Eight would read post metadata and comments in an attempt to learn more; six had uploaded undescribed photos they came across on social media to a service that would generate a description with AI (such as Seeing AI); five asked people they knew; two asked human volunteers or employees of visual interpreting services (such as Aira or Be My Eyes [65]); eight had asked the poster to describe the image in a comment, and five participants used residual vision to look at photos. In absence of an image description, participants sometimes liked or commented if the photos were posted by people they care about (eight participants) or if the information they could gather spiked their interest (nine participants). However, 24 participants frequently ignored undescribed images. Interestingly, four participants remarked that the onslaught of undescribed images on their social

Self-Reported Race	#
Black	3
African American	1
Asian American	3
Latinx	3
mixed race Latinx	2
Middle Eastern	2
Indian	2
Native American and Hispanic	1
Ashkenazi Jewish	1
white mixed race	1
white	6
Total	25

Table 1. Self-reported race demographics for the study participants.

media contributed to negative mental health; they generally did a lot of advocacy to make known to their friends that image descriptions benefit them and other blind people, and they felt ignored by this apparent dismissal.

Image descriptions were particularly important to a subset of nine participants, many of whom used social media to advocate for descriptions or worked in the field of accessibility professionally. These participants sought out groups and pages (Facebook) and accounts (Instagram) to curate a more accessible social media browsing experience. These accessible outlets included groups sharing described memes on Facebook and the #ImageDescriptions and #AltText hashtags on Instagram. Two of these participants created separate Facebook groups to facilitate mutual aid in the composition of image descriptions. This labor and advocacy was characterized by recognition of a need for image description education and equitable expectations; they found that their contacts did not know how to write quality image descriptions and that blindness was not the only barrier to composing them. Taylor described the purpose of their group, “*Anybody can join and can use the hearing or the sight they do have to provide image descriptions and transcripts and captions for Deaf or blind or DeafBlind folks. And also for folks who have disabilities, where they can hear and see things, but they can't type a lot or they have brain injury and it's hard to work images into words.*” For these participants, curating accessible feeds and promoting education in image description composition contributed to their overall enjoyment on social media.

Though they post photos much less often than they come across them, 22 participants reported uploading images to social media at least once per month. Twenty-one of these participants reported describing images they post, and most used multiple strategies at once to compose descriptions; 15 attempted to write a description themselves; 13 asked for assistance from a person they know, and six asked volunteers or employees of visual interpreting services.

Given the recent proliferation of personalized, animated caricatures such as avatars and Memojis [5, 45], we asked participants to share their experiences creating one. The 11 participants who had (10 Memoji and 1 Facebook avatar), expressed a mixture of enthusiasm and confusion. The level of detail participants received about others’ caricatures encountered with their screen readers left much to be desired. However, several participants praised the Memoji creation experience for being technically accessible with Apple’s screen reader, Voiceover, and for offering a variety of hairstyles and customizable skin tones. Some participants were overwhelmed with options, sharing that their understandings of visual phenomena were insufficient to make final decisions about their caricatures, which they intended to look like them. Three participants abandoned their initial desires to make Memoji look like them, instead opting to create intentionally non-human-like caricatures. But others chose to trust the descriptions and attempted creation with varying degrees of confidence. Kai explained her experience, “*When they [emoji] are doing the skin tones, they might have three*

Self-Reported Gender	#
non-binary	5
trans man	3
agender	1
trans woman	1
transfeminine	1
cis man	7
cis woman	7
Total	25

Table 2. Self-reported gender demographics for the study participants.

or four people of color. And it's like, no, I don't think I'm that one, that one or that one. So which one do I choose? Whereas with the Memojis, it felt like there were lots of different ones. And once you picked one, there was a little adjusting thing where you could make it lighter, make it darker or more cool, more red undertones and things like that. I guess I may have cool undertones, but it was a percentage thing and I'm like, okay, I don't know how cool they are. It was nice to have options, but as a blind person, it was just a lot to choose from." Marvelette, in contrast, felt confident choosing features that reflected the language she used to describe herself, "*There were so many ways of describing skin tone. And there were so many ways of describing hairstyles and eye color. And this is where I have to trust technology that it is doing what it's saying on the box, but I was really happy to be able to make something that I think kind of looks like me.*" Finally, Aqua mentioned spending a long time composing their Memoji, Googling terms with which they were unfamiliar and paying attention to their connections' Memoji's. They reflected, "*It feels almost detective-like, are you [contact] using this [feature] for the same thing I would use it for?*" Sometimes, as in Aqua's experience, participants mapped what they knew about people to the features they heard described when focusing Voiceover on their connections' Memojis. Others relied on descriptions of their own appearance people had told them throughout their lives and stereotypes they had read about what people of their race and gender look like. Hoang explained, "*There's this understanding of possible characteristics that I might have based not necessarily on me seeing myself, but people's ideas of what I look like and the way that they talk about Asian people in general.*" In summary, several participants were interested in building caricatures that looked like them, taking notice of the representation offered. However, even with general understandings of their own appearance, and technically accessible caricature creation tools, some participants were not confident in articulating their own appearance and interpreting the nonvisual descriptions available.

5.2 Appearance and Identity Descriptions: Behaviors and Preferences

This section concerns participants' behaviors and preferences around describing themselves and experiences being misrepresented, and in what situations they would like to learn others' appearances.

5.2.1 Self Descriptions of Identity. We asked participants to discuss their race, gender, and disabilities, and how they impact self descriptions in different contexts. Participants adopted labels for their identities which reflected their alignment with political positions of that identity. They also shared how they code switch—shifting language and presentation based on context [110]. Participants recognized the value of language, particularly that which would keep them safest, but they also found language limiting, adding context to their chosen labels during the interview to clarify their identities and to explain how their identity evolved over time.

Several participants changed language if they perceived they were in mostly white, cisgender, nondisabled, and professional company. This primarily comprised a switch to phrases they perceived to be more palatable to privileged audiences, such as "person of color" from language they used in company with their communities such as, "brown" or "Black." Some non-binary and transgender participants described their gender in one word, such as "non-binary," rather than elaborating, or they opted to not disclose their gender at all. We expand on participants' disability to exemplify the politics of identity label choices and how code switching was employed.

How to talk about disability has long been discussed in both academic and activist spaces. Some disabled people claim their disabilities before their personhood to assert disability is a legitimate human experience inseparable from the rest of their identity (called identity first language). Others take the dehumanization of people with disabilities as a sign to assert their humanity first (called person first language) [78]. Several participants had strong feelings about how they described their disabilities, and they reflected how these disabilities fit into their identity and how their language

choices aligned themselves with different activist stances. We noticed however, exceptions made for situations when participants interacted with people with disabilities, or those close to them, who may be new to disability. Kai explained both her pride in a disability identity and interest to carefully bring others in, which she did by code switching. “*So blind is something that I identify as, but I also leave space. Cause I get that folks [with visual impairments] aren’t always comfortable with identifying themselves as blind.*” Taylor’s language changed depending on what disability they were referring to by shifting between person and identity first language, and only labeling some disabling experiences as disabilities, “*I’m proud to be blind and autistic, but my chronic illnesses are definitely not a point of pride and I would accept a cure in a second. So I actually talk differently about each disability for that reason.*”

Though rare, some participants specifically altered their language when speaking with blind people. They believed that language, which may be a primary modality through which blind people learn about appearance, does not capture the actual variety of skin tones, for example. Marvelette explained how descriptions of herself changed before blind audiences, “*I have met white blind people that don’t understand that Brown comes in a lot of shades. I don’t think that people really have those more subtle conversations around [race] with blind people in general. I would probably tell [blind people] that I’m Latina, I have caramel, Brown skin because Brown skin comes in a bazillion shades.*”

If language could liberate participants in some instances, such as Marvelette’s detailed descriptions of her skin tone to blind audiences, language also seemed insufficient to summarize identities. Participants told stories about transitioning genders, tracing their heritage, and making sense of their vision impairment, exemplifying moments that depicted the complexities of their identities phrases could not capture. Danielle, for example, is a transgender woman but is still exploring and does not know if that is “*where I will land yet,*” and Creo has yet tracked down their Native American heritage contributing to reluctance claiming it as part of their identity. These examples showed that participants had difficulty expressing their identities with just a few words, words that could become the basis for describing appearance in image descriptions.

5.2.2 Misrepresentation by Others. As participants spoke at length about their appearance and identities, they also shared numerous stories of their races and genders being misrepresented by others. Though participants were asked if they had been misrepresented generally, few shared stories of misrepresentation related to disability. We note that if disability misrepresentation was not discussed at length during these interviews, it is rampant and discussed in other literature [31]. Nineteen participants were aware of being misrepresented and 13 of these believed this misrepresentation significantly negatively impacted them either in the past or currently. Nine of the 11 non-binary and transgender participants found misrepresentation particularly negatively impactful whereas four of the 17 BIPOC participants did so. The remaining six participants who recalled experiences being misrepresented (two non-binary and transgender participants and four BIPOC participants) told us that misrepresentation did not significantly impact them.

During the interviews, we learned about two types of misrepresentation; non-binary and transgender participants were frequently misgendered, particularly when and/or if they did not present as a binary gender, and BIPOC participants’ races were consistently guessed wrong. Aqua is very impacted by misrepresentation as it frequently happens in a context where they have done a lot of preventative labor and advocacy, “*I’ve been described incorrectly before, and it’s really off putting, because I have put so much time and effort into informing people about how to make image descriptions. And so for people to use that as a way to misgender me is pretty painful.*” Sophie shared that she is regularly assumed to be white in blind communities, “*When they find out that I’m Indian, they’re like, Oh, I thought you were blonde or I thought you were a white girl. You talk like a white girl. It bothers me in the sense of like, what am I supposed to sound like?*” One participant changed their name to one rooted in their culture in hopes they would not be assumed white

during phone conversations or by blind people. Others, like Wiles, believed their outward appearance sufficiently coded how they wanted to be represented, and were perplexed at how to respond to erasure, “*You put a photo of yourself or just your body out there in public. And people look at you and see someone who you’re not. And that is a very disturbing feeling. Even though I’m not out, I try to code visibly queer. There’s sometimes that feeling of despair of like, how do I ever get society to recognize me for who I am?*” With their disparate impact on participants, we learned that images, and their descriptions, could be the sites of incorrect or offensive interpretations. Subsequently, their experiences with misrepresentation informed participant image description language preferences, elaborated later.

5.2.3 Self-Reported Image Descriptions. Turning back to social media, this section summarizes what types of information participants included when describing images of themselves. We asked participants to locate a few different photos of themselves that they may post in different contexts and to describe them during the interview. We offered posting context and photo type suggestions including a professional photo, a casual pose, or an action shot. Twenty-two participants located photos; the remaining three did not have photos available to share. Further, eight participants shared image descriptions, but opted to not send the researchers their photos to review as reference points for privacy reasons. We categorized these self-reported descriptions below. Of note is that although we asked participants to describe photos they may post in different contexts, all 22 participants who post photos at least monthly assumed most viewers were people they already knew.

Twenty-one of 22 participants who post photos on social media at least once per month include image descriptions with their posts. The following summary of image description features come from these participants only. Action and clothing were the most popular features of image descriptions, regularly described by 20 participants each with the caveat that participants generally prioritized describing actions over clothing, preferring to share more details about clothing in posed shots. Sixteen and 14 participants, respectively, regularly described an image’s background and accessories including jewelry, hats, glasses, and handbags. Seventeen participants described disability-related access technologies, such as white canes and guide dogs. Only three participants explicitly shared that they aimed to make their image descriptions as minimal as possible. We note that participants most often posted photos on Facebook and Instagram, where no character limit dictates image description length. Kai summarized her typical image descriptions, “*I’ll talk about what the setting is. So if I’m at a park or I’m in my house. And then what I’m wearing. If I’m wearing jewelry, what colors the clothes are. I’ll talk about whatever pose I’m in, if I’m smiling, if I’m looking away from the camera, looking at the camera, just kind of try to be as vivid as possible.*” Sophie, in contrast, was among those preferring minimal information, named the people in images and described the background in one sentence.

Most participants, like Kimsan, did not share information about their appearance to known audiences because they believed, “*My friends all mostly know what I look like.*” This assumption of prior knowledge, to most participants, rendered providing information about their race, gender, and disabilities irrelevant in image descriptions of themselves posted to known audiences.

Only three participants regularly shared their race, gender, and disabilities in image descriptions to known audiences. However, all three do not describe them in every image, instead choosing to remind their known viewers (as they do not post to unknown viewers), of their appearance every now and then as Aqua does, “*I will occasionally describe I’m a light skinned Latinx person with dark Brown curly hair that’s this long. I definitely will describe when aspects of disability are obvious in a photo so that people know those things about me. I don’t add that all the time. I’m like it’s been three photos since I last told you that this is what I look like. The point is that I’m wearing this bracelet.*” Three separate participants clarified that they do not describe their appearance on social media but that they provide a verbal image description of

themselves before speaking to audiences at disability-centered events. To these six participants, they considered this information part of providing basic access to blind people, which could also help to push back on what they perceived were pervasive assumptions by people that unless told otherwise, people are white, cisgender, and nondisabled.

To the 19 participants who post photos at least once per month but who do not describe their appearance to known audiences, we asked in what contexts, if any, they would consider describing it. Eleven speculated they would consider providing an appearance description in images if they knew they would be posted publicly, and 16 speculated they would consider doing so if their image would be posted in the context that they were speaking about their race, gender, or disabilities. For example, Sasqya wanted people who had not met her to understand the importance of her race, gender, and disabilities to her appearance and identity, “*If it was more of a public facing thing It's important to me that they [organizers] highlight what type of cane I was using, what color dress I was wearing, what color skin I had, the type of hair. There's only a finite amount of time they have to describe, but these are important.*” Marvelette explained that positioning herself in discussions of identity superseded whether the audience was blind, “*The experience I have most as of right now is discussing race and in those cases, I do think that is really important to say how you identify. And honestly, not just for me as a blind person, but skin tone and racial identity don't always equate. So I think it's really important in those kinds of conversations to say, I identify as X, Y, and Z. And this is my perspective from where I identify.*” If describing appearance was not common practice on social media, most participants expressed that some situations warranted including it in their self-description.

5.2.4 Interest to Know Others' Appearance. We asked participants to share situations in which they seek out or want to know the appearance of people in photos they encounter on social media or in person. Upon sharing a context, participants were subsequently asked to share the strategies they use to discern appearance information, and challenges to obtaining this information. Overall, participants wanted access to appearance information much more often than they provided in their own image descriptions.

Eleven participants often wondered about people’s appearance and advocated that this information should be made available to them if desired, though except for one, they generally felt embarrassed to ask. They justified that anything available to a sighted person should be made so for blind people, while acknowledging that this access should not come at the expense of hurtful descriptions. They contended that this information could contribute to ongoing awareness of important topics like the ways race, gender, and disability are coded and politicized. Tracey explained, “*If somebody [sighted] sees a photo and has some kind of clues, we should be given comparable information. If you don't provide that, it does feed into, maybe indirectly, blind people don't need this information or they don't make judgements based on this information. If everybody else has access and the goal is to have access to photos through descriptions, then that should be provided, but we've gotta find a way to provide it in a way that's not overly prescriptive.*”

We synthesized four contexts when several participants interacted with people whose appearance was unfamiliar to them and in which they communicated that such information about others was particularly important. These contexts included (followed by the number of participants):

- (1) When identity is the topic of conversation (19).
- (2) To learn representation in media (13).
- (3) To read a room to know their audience and find community with people like them (12).
- (4) When seeking specific perspectives (10).

Participants employed several strategies to obtain appearance and identity information in these contexts. But they clarified these tactics can be time consuming, are often based on stereotypes, and may still be insufficient.

Identity as conversation topic. When identity was either the explicit or implicit topic of conversation, 19 participants found that appearance information about those involved was extremely important. Explicit conversations about identity may involve those on race, gender, or disability. Implicit conversations could involve topics about issues, such as police encounters or housing justice, that disproportionately impact marginalized people, or those where group language is used. For example, Emmett explained, “*I tend to listen more to people who have had the experience of not being white or not being cisgender or being disabled and chats. And since I can't really see everybody who is participating in this dialogue, I find it a little bit more important to figure out where others are coming from.*” To Emmett and others, appearance seemed to give sighted people a clue as to how to interpret commentary on identity, and in its absence, they argued that disclosure of such information was necessary for them to have an equal opportunity to understand what was being shared. Sometimes, not knowing this information created conflict, as with a Black DeafBlind cis man participant, Blue’s, experience, “*There was a Black activist that I followed. We got into a disagreement about something that she said. And I don't know the arguments that got her telling me that I was being typical for a man communicating to someone who was presenting as a woman. In my head I was thinking, Hey, I didn't even know if you were a woman. I hadn't even really assigned a gender to you.*” Finally, participants also wanted appearance information when they heard in group language. Jay exemplified the complexity of nonvisually detecting appropriate language use by recounting a conversation with a sighted friend, “[Sighted friend] said that somebody had posted [in group language]. And I did not know that the person who posted was a person of color, but [sighted friend] really quickly googled, this is the use of this word. And based on that, he was like, Oh, the person who's using it, that's fine. This is an interesting piece of information because had I not known that [person's appearance] and had I seen somebody use that word, I would probably challenge them on it.”

A consequence of this uncertainty was that several participants to lurk, rather than participate, in identity-based discussions on social media so they avoid situations like Blue’s. However, as moderator of a Facebook group, Aqua has intervened if they perceive white group members may be commenting on posts where they have been requested not to. Aqua relies on qualifying language to position themselves to complete their moderator responsibilities as a blind person by writing something like the following on a comments thread they perceive needs moderation, “*I don't know this, but if you are a white person, as a moderator I'm telling you this is not somewhere you should be.*” Though many avoided engaging in discussions about identity as to not make mistakes about those with whom they were conversing, some participants got into confusing and negative interactions online when they were assumed to be able to see profile pictures, and others, like Aqua, were careful with their language to reiterate group rules while clarifying they could not ultimately judge the appropriateness of behavior as they could not discern someone’s appearance. In this case, insufficient image descriptions could prevent participants from engaging in discussions which may assist in education and understanding about race, gender, and disability.

Reading a room. Second, some participants (13) wanted to be able to ‘read a room’ to understand when and how to code switch and to find community with people who appeared similarly to them. For example, Yvonne chooses what and how much to share more confidently when she knows appearance information. “*We're about to talk about race and is the person that I'm talking to a person of color? And of course that hits on all kinds of stereotypes, but sometimes I want to know, is this a person who may have a lived experience, or is this person more than likely an ally? Cause sometimes that shapes how vulnerable I am with people.*” Similarly, these participants were constantly interested in building community and sometimes lamented when they learned information after knowing someone for a while because of opportunities they may have had to relate to them earlier in their relationship. Sophie shared a relevant experience, “*One coworker was talking, I learned that I was making way less than this guy. And of course, the Black woman, making way less. And I had no idea she was Black and I was like, Oh, I might have had these types of conversations with her previously if I knew.*

Cause its part of my race being Brown." Participants made assumptions nonvisually about who people were by judging whether people's voices had characteristics stereotyped as belonging to people of specific races or genders, and listening carefully to what language was used (for example, if someone says 'we' when talking about trans community). Finally, some participants, like Kai, relied on others reaching out to her, something she and others found disappointing but more reliable, "*We all had to post a picture to introduce us to this really big, large community group. And Then I got an invite to this Black Indigenous people of color group for the training. So I didn't really have to do that work. They saw my photo and then they could instantly ask me, but I didn't unfortunately have that same opportunity.*" To these participants, building community often led them to rely on stereotypical identity presentations and verbal identity disclosure. Such reliance meant they sometimes missed opportunities to connect with others like them. Otherwise, they hoped others who looked like them would take the initiative to reach out.

Representation in media. A third context when participants (12) wanted to learn appearance was to learn representation in the media. Representation in the media played an important role in giving participants a status on how representation elsewhere may be, to help facilitate conversations about identity and politics, and to help blind participants to identify products and public figures they'd like to purchase and follow, respectively. Kai explained its importance, "*If you're only used to seeing a certain group of people and you don't get [those descriptions] on social media, you're going to assume [the whole world] is pretty much the world around you.*" Knowing this information could also help facilitate conversations, like those had with children as Hoang explained, "*Any time we watch media, we [Hoang and children] talk about it. So there's a lot of discussions about race, gender, disabilities. We spark conversations about how these characteristics are being depicted, and whether or not they're correct or their hurtful. I watch using the audio description track before I show the kids the movie. That way I have a better understanding of possible ways to describe certain scenes or what they might potentially be seeing. And then I can interact with them based on the information that I have through the audio descriptions. But there are definitely instances where audio description is vague or maybe using stereotypical language. So for the most part, I don't currently have a way of verifying whether or not that information is correct.*" Additionally, learning representation in the media led participants to judge whether products suited them and to follow certain public figures. This was pertinent for getting clues about how to achieve a desired gender presentation and which products worked for their body and hair types. Tracey clarified, "*Having never been able to see, I'm not seeing how people look passively as I go through life. I was socialized as a [misgender], so there's a lot of [differently gendered] fashion stuff that I just don't know.*" Some participants mentioned reading model height and the size clothes they were wearing as ways to gather information, but most were still unsure whether products were modeled on people with body types, skin tone, and hair similar to theirs. Some participants, including Sasqya, found this information not just helpful, but crucial to making purchasing decisions, "*My hair is wavy and frizzy and needs maintaining. So I look for specific products. And if it's being catered to somebody who doesn't look like me or doesn't have my hair, I don't trust it as much, even if it says all the right things.*" Insufficient product model descriptions led these participants to seek more human assistance than they would prefer to engage while shopping. Finally, participants found enjoyment and empowerment witnessing people like them in the media. Sasqya explained that not growing up "*seeing people like me*" meant she was constantly wondering whether characters would resemble her, and Jay felt fulfilled "*To see people with disabilities engaging in a variety of activities. To clarify, it's not for inspiration or anything. It's just satisfying to see that somebody is also having a good life.*" To understand representation in the media nonvisually, participants turned to image and audio descriptions which as Hoang explained, did not consistently provide appearance information, or they searched for text-based media written on the topic. This meant that what they learned about appearance information about media figures was reduced to what primarily sighted commentators found important to share.

Seeking particular perspectives. Some participants (10) also thought appearance information would be helpful when they seek particular perspectives. They found that photos were often assumed to do this work by for example, depicting an author they were considering reading. They shared that despite the increase of curated recommendations based on identity, finding this information in text was not always easy. Kai shared her experience seeking a very specific product from a Black-owned business, “*I found all these [products], but I had to do a lot of research to find out if this [owner] was actually a Black woman. I had to really go to different blogs and different websites and then look the person up and other work they've done. It just felt like it took a really long time.*” Another workaround occurred through strategic conversation, as to not ask directly, but in hopes of getting the sought answer, as Bassam did when he discovered new leadership at an Asian-owned business he frequented, “*I brought it up in a roundabout way. I was like, ‘didn't the previous owners used to be a Korean family?’ And [owner's] like, ‘yeah, I'm Korean too.’*” Participants acknowledged resources for finding perspectives exist in popular media, such as Netflix’s Black Lives Matter recommendations [79]. But they were inconsistently available and finding the information written in text or cleverly asking for disclosure was more time consuming than glancing at photos.

Concerns with too much description. Though participants wanted to learn more appearance information from image descriptions, several were concerned that incorporating appearance into all image descriptions could overemphasize it. Some participants contended that if race, gender, and disability were consistently assumed from appearance, they were done so with uneven rates of awareness; sometimes a viewer may pay attention to other aspects of the image, and to these participants, image descriptions that included appearance information by default could more frequently focus the reader’s attention on appearance than what immediately caught sighted people’s attention. Sophie explained her concern with overemphasis, “*There's messages people can take from writing. What they [image describers] choose to emphasize might bias people who are reading it, and you're not getting the same information as if you were just looking.*” Other participants, like Marvelette, were concerned image descriptions of race, gender, and disability would embolden blind people’s racism, transphobia, and ableism, all while acknowledging that blind people shouldn’t be on principle prevented from knowing. Finally, some participants like Yvonne were concerned disclosing appearance could become a slippery slope that could put undue pressure on minoritized people to share more about themselves than they wanted to. To Yvonne and others, this expectation of disclosure seemed unjust as such descriptions have become important to understand certain situations precisely because of racism, transphobia, and ableism for which minoritized photographees and describers were not responsible. We found this discomfort with recommendations to describe appearance all the time interesting, raising the importance of contextualizing best practices not only for relevance purposes but in protection of those who may feel the most undo pressure to share.

5.3 How to Describe Appearance

Next, we asked participants to reflect on their preferences for how image descriptions of themselves should be composed. Throughout the interviews, whether a describer could contact a photographed person directly or not remained an important distinction, so we divide preferences along these lines.

Participants always wanted to be described with the language they used to talk about themselves. They argued that out of respect, image describers who know the photographed person should consult the person or reputable sources for information on how to describe their appearance. For example, participants believed that using someone’s gender listed on a social media profile was acceptable in lieu of asking the person directly. This preference remained for both close contacts of participants as well as situations, like Sasqya’s participation in a professional event as mentioned earlier, when a content creator may have the ability to contact the photographed person.

Next, we articulate preferred appearance descriptions when the image describer cannot confirm the identity of the photographee or when the photographee does not want all aspects of their identity disclosed to the viewing audience. Twenty-four participants shared that language that described concrete, visual details (e.g., skin tone) was more appropriate than language that represents identities (e.g., race labels). However, participants warned that language changes, so best practices must be regularly updated. Taylor and Aqua who moderate the afore-mentioned Facebook groups assisting users to compose image descriptions summarized their recommendations, which aligned with participant preferences. Approximating skin tone and describing hairstyles may help describers and viewers avoid assuming race; describing clothing, accessories and hairstyles can help describers and viewers avoid gender assumptions, and describing access technologies can help describers and viewers avoid assuming disability.

However, participants brought up complexities of describing appearance. A few non-binary and transgender participants wanted their gender to be correctly assumed in image descriptions as part of affirmation. Some BIPOC participants were concerned that language to briefly describe skin tone feeds into colorism (when using light and dark descriptors) [57] and such phrases inadequately convey the variety of skin tones and physical features. As Marvelette incorporated more details of her skin tone when speaking with blind people, Hoang was also concerned this complexity would be lost in brevity, though participants did not have suggestions for how to enrich these descriptions when the photographed person was unknown to the describer. Hoang explained the limitations of skin tone approximations as to avoid identity assumptions, “*Skin tone, there’s such a range. And I haven’t run into a good image description that included skin tone information in a way that I liked it. I see it as useful information, but I also see it as ambiguous. So medium skin tone. If you don’t have color perception, what does that mean?*” Finally, two participants believed that describing appearance without assuming identity diverged from popular language conventions and withheld information from blind people, believing that if a sighted image describer made assumptions about identity, they should communicate those. Although participants were generally in favor of appearance being described in concrete, less politicized terms, they recognized limitations and complexities of this careful language. These complex viewpoints raised the importance of maintaining dialogue on respectful language as to not lose track of evolving and varied preferences.

5.4 Preferences of AI-Generated Image Descriptions

Most participants were familiar with AI-generated image descriptions provided in their iOS Photos library, on Facebook, and through the Seeing AI app. We note that to our knowledge, most AI-generated image descriptions do not communicate race, gender, or disability. However, Seeing AI sometimes provides an age and gender description after a user uploads an image to the Scene Descriptions channel of the app. Participants generally did not trust these AI-generated image descriptions, claiming that they would seek human assistance if an image’s contents was important. They reported using the features nonetheless for purposes of reading text in images, deciding whether to make the effort to learn more about a photo, or as a form of entertainment. A few participants, like Hoang, appreciated the clues automatic alt text could give when checking in on people important to him, “*Sometimes hearing that information will cause me to dig into the thread to see whether or not it’s more interesting and to not bypass it altogether.*” However, these same clues were frustrating to some who felt the information was insufficient on its own and most prominently reminded them of access they were missing, as Taylor characterized, “[AI-generated image descriptions] are such a teaser, it’s just enough information to know what I’m missing, but not enough information to actually know what’s happening.”

Many participants referenced Seeing AI’s Scene Description channel as a form of entertainment, using it with groups of friends to laugh about inconsistent and often inaccurate gender and age judgments. For example, a cis woman participant, Kimsan, has allowed her child and friends to play with Seeing AI, ‘*I’ll say something like, ‘Oh, it said*

'you look like you're 16,' and then the girls get all excited and happy. So I think it's more for fun and games, but never something to take seriously." But misrepresentation also negatively impacted several, particularly those who were non-binary and transgender. Some were harmed by the act of misidentification and some were frustrated that gender judgements seemed to be based on a binary gender classification. Creo, who is neither a man or a woman as Seeing AI guessed, found it entertaining but also frustrating, "*Just like humans, [Seeing AI] typically tries to shove you into one characterization or another. I've taken pictures of myself just out of sheer curiosity. And sometimes I'm a 35-year-old woman looking happy. Sometimes I'm a 50-year-old man looking happy. ... So it seems pretty confused about my gender and my age.*" Emmett situated this act in the classification system that created it, "*It's definitely frustrating having this sort of technology get integral parts of my identity wrong. And I find it frustrating that these sorts of apps only tend to recognize two binary genders.*"

After sharing their current experiences with AI-generated image descriptions, we asked participants to share their perspectives on future AI-generated image descriptions that may describe photographed people's race, gender, and disabilities. Most participants were excited about increased access to appearance information in image descriptions and considered AI a promising avenue to continue developing, with priority to improving descriptions to be accurate and respectful. Parker summarized this perspective, "*I favor something rather than nothing. Even if the identities of some people that it describes are wrong.*" However, this enthusiasm came with a lot of questions and concerns. Prior to the interviewer bringing up AI bias, 15 participants mentioned known instances including its misuse by law enforcement, identity-based discrepancies in medical care triaged by AI, automatic job application review that rank minoritized candidates lower, and the racist history of photography which is tuned to capture lighter skin tones.

As conversations evolved to explicitly weigh accessibility and the potential for misrepresentation, participants diverged in their preferences. Six continued advocating the release of AI-generated descriptions of appearance, hoping they would improve to address bias but considering the access they may provide as important to weigh with their perspective. nine participants speculated some design considerations that could ease their concerns about AI and bias and remained hopeful that if addressed, they could benefit blind people. Ten participants remained skeptical that AI-generated descriptions were appropriate given current patterns of discrimination in greater society.

Participants ideated considerations that may ease their concerns about bias seeping into AI-generated image descriptions of appearance. First, participants questioned what companies were doing to insure features' potential impacts were evaluated. As part of this sought accountability, some, including Rico, stated requests such as, "*It'd be kind of nice to know that there's a diverse group working on this and not predominantly one type of person.*" Second, 24 participants argued AI-generated descriptions would be more respectful if they used appearance rather than identity presumptive language (as discussed earlier). Third, 24 participants believed for AI-generated descriptions to be ethical, it was imperative that people have control over how they are described and the power to opt in/out of being analyzed for the purpose of generating image descriptions. Suggestions included opting-in to a feature where a social media platform would incorporate disclosed profile information, such as pronouns, into image descriptions or allow screen reader users to modulate the amount of information they hear. Additionally, echoing Sophie's concerns that default appearance descriptions may water down their importance, participants like Blue wanted control of which contexts garnered appearance descriptions. To participants, these controls may help those photographed be more aware of how they were represented in descriptions and reduce nonconsensual analysis and possible misrepresentation. Still, participants had some concerns about the design of such supports; those who had not disclosed their gender publicly wanted privacy controls to customize which people could read which types of image descriptions, for example.

Another subset of participants (9) assessed potential harms of AI describing appearance high enough that they recommended features should not be released now, but their hope for a future release remained. This preference was underpinned by four reasons; first, people of minoritized races, genders, and disabilities already experience discrimination; they did not want AI to exacerbate this. Tracey anticipated a future of being regularly misgendered by AI, “*It’s just one more microaggression that I have to put up with from technology that’s supposed to help. I would hate if every time I upload a photo [to Facebook], I have to change it or do something so that it doesn’t say [age, misgender]. ... I understand the argument, I’m sure many people say that some description is better than none. Well what if part of the picture is to illustrate to the viewer that, Hey, I am trans, you know, I may have been [misgender] assigned at birth, but I am not.*” Second, participants believed biased AI would perpetuate bias in the blind community, potentially with no recourse or incentive for screen reader users to check the information given them if AI descriptions are their only information source about an image. Third, some participants felt that if released features caused documented harm, this sent a message that the dignity of people with disabilities was disregarded by having lower standards for the ethical review of access technologies. Yvonne explained, “*The people that are impacted are people that are often already bearing the brunt of so much other stuff. And all of the anti-Black messaging that I’m seeing from all kinds of people on Facebook right now would just be compounded by the fact that AI is giving me biased information. If you are doing something where you are filling in people’s perceptions, because they are not perceiving with their sight, which is such an overriding sense and you’re filling it in with something that is biased like that, Ooh, that just seems so dangerous to blind people. And then there is just so much wrong with the idea of okay, it is biased, let’s just give it to people anyway.*” Finally, Sincere highlighted a fourth reason why deployment should be withheld for now: the potential harm of having to correct wrong information. She explained, “*It’s [AI image description] our first visual image of the photo. So if it’s giving us incorrect information to start off with, that’s something that we have to process. Now, once we process that and we find out that it’s incorrect because we have someone sighted and looking at it, now we have all this comparison going on in our mind.*” These comparisons and the unlearning they would necessitate could waste time and cause even more harm to understandings of race, gender, and disability. These concerns led this subset of participants to advocate for more research and review before AI describes appearance.

Finally, 10 participants remained skeptical of AI’s capabilities and doubted whether recovery of existing harms was possible given pervasive exploitation of user data by companies providing such services, racism, transphobia, and ableism. First, some participants, like Hoang, were concerned that collection and dissemination of personal information, such as appearance descriptions, could be exploited in nontransparent data flows. These participants made clear that the idea of AI in a vacuum was not problematic, but they were concerned about its prior record of causing harm to people like them and its reliance on static categories which cannot encapsulate human complexity. Taylor summarized this challenge, “*I’m a little skeptical how helpful AI can even be because there is so much about identity that can’t be broken down into these discrete units. Two people who look the same might have different genders, races, and disability statuses. An AI would not be able to know that.*” These participants cited pervasive discrimination, previous instances of data exploitation, and the limitations of AI’s reliance on discretely categorizing people as reasons to caution applying AI to describe people’s appearance altogether.

6 DISCUSSION

From our interviews, we learned how a subset of screen reader users who are BIPOC, non-binary, and transgender understand human appearance in photographs. In discussing appearance, AI, and bias, we discovered a tension between wanting to know others’ appearance, vulnerabilities of expecting such information be provided nonvisually, and



Fig. 1. A potential description for this image may be: “[A Black, disabled, non-binary person] with a filtering face mask walks down a neighborhood street with one hand in their pocket and the other hand on their cane. They have a short mohawk and are wearing a jacket, shorts, tennis shoes, and glasses.” The description for this image might vary based on audience and knowledge of the photographee. The identity phrase (in brackets) includes both the photographee’s preferred identity description and word ordering (identity first disability language). When identity preferences of the photographee are not known, the identity phrase might be replaced with an appearance phrase such as “a person with darker skin.” When the audience knows the person in the image well, it could be just their name: “Leila.” If the image is used to promote popular media such as music and film, both the name and identity/appearance description may be included. [Source: Disabled and Here (CC-BY)]

concerns of how AI may increase access at scale with risk of perpetuating bias. In what follows, we engage our findings on (non)visual sensemaking and digital media, appearance and image descriptions, and the implications of AI generating them.

6.1 (Non)Visual Sensing of Appearance and Digital Communications

As previous research has found, negotiating appearance does not evade blind people [33, 63, 81]. Our participants were interested in knowing the appearance of others, so they employed a variety of nonvisual strategies to learn the appearance of people of interest. From these findings, we commentate on HCI research to promote justice and the impact of our design and research priorities in cementing inaccurate understandings of how race, gender, and disability are negotiated. First, as scholars contend that HCI must reckon with the racism, ableism, and transphobia we perpetuate both within and outside the field [55, 82, 103], we anticipate growing interventions toward justice. Part of this work will concern education and awareness about the ways we exude prejudice and necessitate unlearning and

reparation. To provide holistic representation, this education will need to be accessible in addition to delving into the ample and multisensory ways, not just visual, we make sense of contested identities. We argue that much of this work could amplify what exists; the Crip Camp Virtual Experience [26] and Sins Invalid Performance Project [101] are two examples which unpack systemic harms of ableism, racism, and transphobia in an accessible manner. For example, as Sasqya mentioned, some disability-centered events increase nonvisual access to their presentations by requesting presenters to describe their visual appearance at their remarks' outset. The design of our awareness and education, presentations, and the digital media systems which support our academic work could follow suit.

Further, the politics of appearance as described by our participants also challenge the ways we use images as an object of research. From judging personality traits [56, 100] to diagnosing disability [11], it is not only end users of social media platforms putting stock in what others should know from a profile picture. Our research choices elevate visual dominance further when we argue such imagery can assist with understanding people and making decisions based on those judgments, often done altruistically towards health and public safety. But this work shows how well-meant access without thorough ethical review stands to harm those already impacted. As participants carefully separated language which may describe concrete aspects of their appearance from interpretations of their identities, HCI must increase the care and accountability to modulate the claims we make which are based on image analysis.

Finally, our interviews reveal opportunities to reflect on the ways digital media overemphasizes knowledge acquisition through visual sensemaking. Profile pictures are prominently positioned on social media and users take advantage of this placement to perform identity through the image that supposedly everyone can see [19, 62, 71, 126]. These symbols not only pair with other profile information, which may include nonvisual content, but they travel and contribute to the work of comments, becoming their own sites for appearance and identity negotiation [90]. Our participants, like Blue, experienced miscommunications based on the assumptions that vision is a shared language on social media. Risks like this led some to divest from identity-based discussions with concern they would make inappropriate contributions. Simultaneously, some participants believed that enforcing disclosure of appearance and identity could overemphasize it in favor of the messages they wanted to get across and could put already minoritized people at risk of being outed or harassed. We find opportunity in the design of digital media to promote multiple forms of introductions to diversify the ways users may present themselves. For example, ample virtual meeting recommendations like those in [13] teach attendees to add their pronouns to the end of their last name on their video conferencing profile. But as this hack demonstrates, platforms lack widespread and explicit support for such disclosures, if users desire to provide them. Content creators and moderators may think critically about community standards which ask posters to question and minimize the assumptions they expect readers to make about their identity. An open challenge remains in how to gently nudge users to consider offering nonvisual ways to express their identity while respecting users' choice and authority to disclose their identity. However, carefully considered design decisions like these could expand how people relate in online interactions by promoting nonvisually accessible forms of information sharing while presenting them beneficial to everyone (such as stating pronouns).

6.2 Appearance and Image Descriptions

As with prior work [73, 104, 105], screen reader users have contextualized preferences for what is considered important in image descriptions. But these have not been well incorporated into best practices [116, 117]. While some popularly available best practices acknowledge that describing appearance is sometimes important [21, 67], we outlined six contexts when participants speculated that additional appearance descriptions would particularly help them. These included: (1) avatar creation; (2) encountering unknown people; (3) during discussions about identity and appearance;

(4) when seeking to read a room and find community; (5) learning representation in media; and (6) seeking specific perspectives. Several participants believed that blind people lack awareness and language to talk about prejudice like racism, ableism, and transphobia, which motivated preferences for more appearance information in image descriptions overall. But as previously mentioned, participants were simultaneously concerned about overemphasis and undue burden placed on already minoritized people. Here, we find tensions between practice and preference and access and expectation calling into question what we can ask about appearance descriptions when they remain both uncommon and a sensitive topic. With these tensions in mind, we offer suggestions for image describers and platforms hosting images to move forward.

First, the preference upon which participants largely converged was that the language of appearance versus that which presumes identity is different. For example, appearance descriptions may include skin tone, hairstyles, makeup, accessories, clothing, and access technologies. Identity descriptions, on the other hand would consist of race, gender, and disability labels. But these differences remained varied and situated, often determined by personal preference, what code switching in context may warrant, and what have become popular terms to describe identity rather than any longstanding language conventions. While participants were agnostic about some of their identity labels, other language choices symbolized long and contentious journeys to declare their identities. In these cases, usage of the correct terminology was imperative to their affirmation and wellbeing. As such, participants' strongly advocated appearance language whenever describers cannot know their identity. But they also emphasized image describers should make effort to confirm preferred labels, as well as preferred word ordering (such as in the case of identity and person first disability language preferences) with whom they were describing as much as possible, stay up-to-date on community language preferences, and continually question what they can know from an image and what they are assuming. More research is necessary for working through tensions of image description language as liberating and limiting; whereas skin tone seemed to be acceptable to describe by most participants, some understood the categories could feed into colorism [57, 121] and other forms of discrimination, such as cameras capturing lighter skin tones differently.

Second, the participants who made one appreciated being able to create avatar and memoji caricatures, but they were overwhelmed with several options with insufficient descriptions of the end results. However, avatar creation is bounded in that there are not infinite options for personalizing caricatures. As such, avatar creation may be an educational opportunity for designers to communicate particularly detailed descriptions of appearance.

Third, participants differentiated whether posts were viewable publicly in which case an appearance description may be more appropriate, versus photos shared among known contacts for whom such information may be redundant. Profile pictures, for example, represent a case where appearance information could be encouraged whereas appearance may be recommended de-prioritized in friends-only posts. Additionally, professional content creators, particularly of commercial websites featuring media figures and clothing and accessories have great opportunity to expand appearance descriptions across their public offerings. In cases where including the name of famous people may be perceived as more accurate and concise than an appearance description, we advocate also including both a name and appearance descriptions to assist blind people in understanding representation in the media and in finding sought perspectives.

Finally, image describers and accessibility researchers may apply their skills to raise awareness of identity-based and uneven media representation. Participants missed out on identity-based conversations with concern they were not getting information communicated through commentators' images, and several remarked identity and inequality are more generally under-discussed in an accessible manner in the blind community. As such, we recognize relevant examples to follow including the Talk Description to Me Podcast [69] and the ProtestAccess Twitter account [87]. These efforts use the act of image description to discuss protests and identity-based inequities. Accessibility research could

similarly promote image descriptions and tools that may not only serve a functional purpose, but to similarly raise awareness about under-discussed topics.

6.3 Equity-Informed Future Directions for AI and Image Descriptions

Participants were generally excited about any innovation which would proliferate image descriptions at scale. As recent research has shown [39, 74] visual information is proliferating online and it is not become more screen reader accessible. If AI-generated image descriptions are currently unsatisfactory, participants believed they may still be a promising solution to solve this problem. Participants accepted that technology is not perfect; several claimed they did not trust AI-generated image descriptions because of their inaccuracy, but we note that in practice, screen reader users trusted them more than they should [68]. Additionally, participants had concerns about the potential for harm to compound if screen reader users do not realize they are reading an inaccurate description. Participants also detailed concerns about AI bias; 15 brought it up before the interviewer asked, citing known instances of identity-based differences in the way AI classified people and the resulting inequities. To some, these biases might be remediated through more accountability from producers of these systems to demonstrate commitments to recognizing people equally and not misusing data. Others remained skeptical, believing AI-generated appearance descriptions, at least currently, was an inappropriate application of AI given its recorded misuse and the milieu of injustice in which the technology is situated. Finally, participants talked through this tension of access and ethics, calling into question the types of tradeoffs we ask of end users, particularly those, like our interviewees, already among those most negatively impacted by these innovations.

These interviews lead us, along with others [10, 11, 95, 97, 120] to strongly caution the use of AI that analyzes humans and human data, and we extend this explicitly to AI which generates image descriptions of people. Participants remained optimistic that de-biasing AI was possible and ethical usage may be on the horizon given increasing awareness of AI bias, but its history of misuse is characterized by secrecy and evading accountability [10, 80]. We instead argue that automation could assist human image describers to, for example, suggest what types of content and language, including that which describes appearance, may be appropriate in image descriptions. Such innovations may also gently prompt users to consider what they actually know about those they are describing, and help initiate communication with those being described when possible. Supports could point image describers to instances where such information may be authored by the person being described such as on publicly-available profiles. However, changing identities complicate the utility of these strategies, so care must be taken to keep those being described in control and aware of opportunities to update their appearance and identity description preferences. Additionally, participants also advocated that photographees have power to opt in and out of AI image analysis, to control who may read specific descriptions, and to choose the language AI may use to describe themselves.

Finally, and most troublingly, we raise concern of the position we put end users in when we neglect nuanced ethical review. AI applied for accessibility purposes demonstrates an interesting case for examining power imbalances as a particular feature may provide unparalleled access but which also feeds into systems that further marginalize some of those who may benefit from the feature. Accessibility advocates may argue that deploying accessible solutions as early as possible is the best option. But participants showed that such decisions may lead to features which label people in reductive and inaccurate ways, such as with a binary gender classification. While not consequential to some, binary gender classifications harmed others who were non-binary and transgender. And to these participants, such harm would not be a one-off event but could ruin entire photo libraries and social media experiences. In this case, learning potential harms required casting a larger net than easily-available community partners who may be focused on improving one

axis of inequity, like accessibility. Neglecting to attend to a wide variety of those impacted by the technologies we advocate for risks replicating the injustice inaccessible technology creates in the first place, such as that perpetrated when a company recognizes an initial release may not be accessible while committing to add accessibility features later. Additionally, such neglect may further distance the voices least represented and most needed in these conversations. Technology then deployed to increase access may cultivate dependence by users who may not have another option, making it more difficult for them to resist the harms such features may promote. Indeed, image descriptions are so scant, that several participants conceded that innovations which could scale their presence may be worthwhile even if they could cause harm. With others, we argue that intentionally seeking out groups who have been shown to be most negatively impacted must occur from the outset [88]. We also call for greater attention to the ways access technologies contribute to perpetuating prejudice, and the uptake of philosophies of design wherein if access may benefit a perceived majority but harm end users who are multiply marginalized, these factors are accounted equitably [1, 6, 22, 24, 27, 51]. This means material and public commitments that technology will not be released until it is both informed and evaluated by communities to benefit those most impacted.

More concretely, we might begin by taking up methods which prompt stakeholders to reflect on known potential harms and use tools of media, history, and design to imagine harms which may be yet discovered. Additionally, open-ended ideation may be triangulated with scaffolding. During interviews, we began conversations on AI-generated image descriptions by not bringing up AI bias to give participants an opportunity to share if it seemed relevant. Afterwards we intentionally mentioned AI bias and specific examples, such as the Gender Shades project [17]. Such decisions insert subjectivity into study protocols [8] but we believe acknowledging this subjectivity to have these conversations is essential toward widening the perspectives we can garner from end users. As end users of image descriptions, participants easily connected material examples of what AI-generated image descriptions of race, gender, and disability may be. But they entered with disparate understandings of AI bias ranging from direct experience to being completely unfamiliar. Participant enthusiasm for AI to solve image description inaccessibility displayed at the outset could have easily cushioned technocentric biases by HCI researchers. But asking participants to weight benefits and harms, and materializing what some of those may be in effort to balance the knowledge they had about each was an important step towards nuancing our findings.

7 LIMITATIONS AND FUTURE WORK

While our interviews make necessary inroads to multidimensional ethical review of technical interventions, our qualitative approach offers partial perspectives on the topic. In particular, our convenience recruitment and explicit conversations on race, gender, and disability led us to participants who disproportionately advocated for image descriptions and who were comfortable speaking about their identity. Future research may engage larger audiences. In contrast, our wide net to touch on the variety of perspectives under-represented in this work by recruiting anyone who is blind and BIPOC or not cisgender meant we only provide surface-level descriptions of experiences and concerns. Future research should scope to specific impacted communities to add depth to this research. Additionally, we note that participants had most experience authoring and encountering image descriptions on social media to known audiences, but many of the contexts in which they desired appearance information constituted disparate domains curated by professional content creators. We recommend that future work directly engage professional image describers and content creators to understand barriers and opportunities for expanding image descriptions to address representation in the media. Finally, we recognize that inquiring preferences (such as image descriptions of appearance participants rarely actually wrote or AI-generated image descriptions that yet exist) are speculative. While extreme care should be

taken as these topics are laden with harm, studying reactions to prototypes of potential solutions used as probes will be instructive for learning how concerns will play out in practice.

8 CONCLUSION

In this paper, we engaged 25 screen reader users who are also BIPOC, non-binary, and/or transgender on the topics of appearance, image descriptions, and AI's generation of them. While assumed primarily visual, appearance is a sociomaterial phenomenon which blind people negotiate with nonvisual workarounds and understandings of visual representation. In parallel, image descriptions remain scant online, and AI may offer scalable solutions to proliferate their presence and therefore increase access to visual information by screen reader users. In our interviews, we learned that participants had extensive experiences developing their identity but were regularly misrepresented by others. They engaged numerous strategies for understanding appearance, which was often sought in certain contexts, yet they rarely described their appearance in practice, calling into question tensions of providing access and what may be safe or reasonable to ask people to disclose. Participants viewed descriptions of appearance and identity extremely different and thought it essential for image describers to understand the distinction; when possible, they wanted their preferred language used to describe themselves, but they strongly recommended less politicized and more concrete description of appearance when the photographee's identity could not be confirmed. Participants were excited about the potential for AI to increase access to image descriptions, but they had significant concerns about its accurate and ethical deployment. From these lessons, we recommend increased commitments in HCI to destabilize assumptions that vision is a shared language in communications, image descriptions which directly engage appearance in appropriate contexts, and guidelines which continually probe image describers to evolve their practices with community preferences and question their assumptions when crafting image descriptions. Finally, along with many other scholars concerned with AI ethics, we caution the deployment of AI-generated image descriptions of appearance, finding the risks currently not well enough understood or mitigatable to warrant accepting imperfect technology. Instead, we recommend that automation turn to support the human authoring of respectful image descriptions, deeper engagement with harms and benefits in our research procedures, and strengthening partnerships among accessibility research and AI activism to take seriously the role community-led accountability may play in creative and ethical applications of AI for accessibility.

ACKNOWLEDGMENTS

We thank all of our participants for sharing their important perspectives. We also thank Mary Fernandez, Rilee Sloan, and Calvin Liang for their contributions. This work was partially supported by the National Science Foundation.

REFERENCES

- [1] Algorithmic Justice League. 2020. *Unmasking AI harms and biases*. <https://www.ajl.org/>
- [2] Bobby Allyn. 2020. ‘The Computer Got It Wrong’: How Facial Recognition Led to False Arrest Of Black Man. <https://www.npr.org/2020/06/24/882683463/the-computer-got-it-wrong-how-facial-recognition-led-to-a-false-arrest-in-michig>
- [3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300233>
- [4] Subini Ancy Annamma, Darrell D Jackson, and Deb Morrison. 2017. Conceptualizing color-evasiveness: using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethnicity and Education* 20, 2 (2017), 147–162. <https://doi.org/10.1080/13613324.2016.1248837>
- [5] Apple Inc. 2020. *How To Make Your Own Memoji*. <https://apps.apple.com/us/story/id1445637997>

- [6] Mariam Asad. 2019. Prefigurative Design as a Method for Research Justice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 200 (Nov. 2019), 18 pages. <https://doi.org/10.1145/3359302>
- [7] Stijn Baert. 2018. Facebook profile picture appearance affects recruiters' first hiring decisions. *New Media & Society* 20, 3 (2018), 1220–1239. <https://doi.org/10.1177/1461444816687294> arXiv:<https://doi.org/10.1177/1461444816687294>
- [8] Shaowen Bardzell. 2010. *Feminist HCI: Taking Stock and Outlining an Agenda for Design*. Association for Computing Machinery, New York, NY, USA, 1301–1310. <https://doi.org/10.1145/1753326.1753521>
- [9] Roy F Baumeister. 1982. A self-presentational view of social phenomena. *Psychological Bulletin* 91, 1 (1982), 3–26. <https://doi.org/10.1037/0033-2950.91.1.3>
- [10] Ruha Benjamin. 2019. Race After Technology: Abolitionist Tools for the New Jim Code. *Social Forces* 98, 4 (12 2019), 1–3. <https://doi.org/10.1093/sf/soz162> arXiv:<https://academic.oup.com/sf/article-pdf/98/4/1/33382045/soz162.pdf>
- [11] Cynthia Bennett and Os Keyes. 2020. What is the Point of Fairness? *Interactions* 27, 3 (April 2020), 35–39. <https://doi.org/10.1145/3386383>
- [12] Cynthia L. Bennett, Jane E. Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173650>
- [13] Berkeley Lab. 2020. *Zoom Tip For Gender Pronouns (April 2020)*. <https://diversity.lbl.gov/2020/04/14/zoom-tip-for-gender-pronouns/>
- [14] Jeffrey P. Bigham, Ryan S. Kaminsky, Richard E. Ladner, Oscar M. Danielsson, and Gordon L. Hempton. 2006. WebInsight: Making Web Images Accessible. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) (Assets '06). Association for Computing Machinery, New York, NY, USA, 181–188. <https://doi.org/10.1145/1168987.1169018>
- [15] Stacy M. Branham, Ali Abdolrahmani, William Easley, Morgan Scheuerman, Erick Ronquillo, and Amy Hurst. 2017. "Is Someone There? Do They Have a Gun": How Visual Information about Others Can Improve Personal Safety Management for Blind Individuals. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) (ASSETS '17). Association for Computing Machinery, New York, NY, USA, 260–269. <https://doi.org/10.1145/3132525.3132534>
- [16] André Brock. 2012. From the Blackhand Side: Twitter as a Cultural Conversation. *Journal of Broadcasting & Electronic Media* 56, 4 (2012), 529–549. <https://doi.org/10.1080/08838151.2012.732147>
- [17] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, New York, NY, USA, 77–91. <http://proceedings.mlr.press/v81/buolamwini18a.html>
- [18] Matthew Carrasco and Andruid Kerne. 2018. Queer Visibility: Supporting LGBT+ Selective Visibility on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173824>
- [19] Jiang Chang, Hailong Ren, and Qiguang Yang. 2018. A virtual gender asylum? The social media profile picture, young Chinese women's self-empowerment, and the emergence of a Chinese digital feminism. *International Journal of Cultural Studies* 21, 3 (2018), 325–340. <https://doi.org/10.1177/1367877916682085>
- [20] Kathy Charmaz. 2014. *Constructing grounded theory*.
- [21] Alex Chen. 2020. *How to write an image description (July 2020)*. <https://uxdesign.cc/how-to-write-an-image-description-2f30d3bf5546>
- [22] Coalition for Critical Technology. 2020. *Abolish the #TechToPrisonPipeline (June 2020)*. <https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16>
- [23] Rachel Cohen-Rottenberg. 2015. *We Need Your Help: Specifying Race and Gender in Image Descriptions (July 2015)*. <https://thebodyisnotanapology.com/magazine/we-need-your-help-specifying-race-and-gender-in-image-descriptions/>
- [24] Sasha Costanza-Chock. 2020. Design Justice: Community-Led Practices to Build the Worlds We Need. <https://doi.org/10.7551/mitpress/12255.001.0001>
- [25] Kimberle Crenshaw. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Policies. *University of Chicago Legal Forum* 1989, 1 (1989), 139–167. <http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8/>
- [26] Crip Camp Impact Campaign. 2020. *Crip Camp: The Official Virtual Experience*. <https://cripcamp.com/officialvirtualexperience/>
- [27] Data for Black Lives. 2020. *Data for Black Lives*. <https://d4bl.org>
- [28] Bryan Dosono and Bryan Semaan. 2018. Identity Work as Deliberation: AAPI Political Discourse in the 2016 US Presidential Election. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3174207>
- [29] Emory James Edwards, Cella Monet Sum, and Stacy M. Branham. 2020. Three Tensions Between Personas and Complex Disability Identities. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3334480.3382931>
- [30] Nirmala Erevelles and Andrea Minear. 2010. Unspeakable Offenses: Untangling Race and Disability in Discourses of Intersectionality. *Journal of Literary & Cultural Disability Studies* (2010). <https://doi.org/10.3828/jlcds.2010.11>
- [31] Heather A. Faustett, Kate E. Ringland, Amanda L. L. Cullen, and Gillian R. Hayes. 2017. (In)Visibility in Disability and Assistive Technology. *ACM Trans. Access. Comput.* 10, 4, Article 14 (Oct. 2017), 17 pages. <https://doi.org/10.1145/3132040>

- [32] Glenn Fleishman. 2000. *Cartoon Captures Spirit of the Internet*. <https://www.nytimes.com/2000/12/14/technology/cartoon-captures-spirit-of-the-internet.html>
- [33] Asia Friedman. 2016. “There are Two People at Work That I’m Fairly Certain are Black”: Uncertainty and Deliberative Thinking in Blind Race Attribution. *The Sociological Quarterly* 57, 3 (2016), 437–461. <https://doi.org/10.1111/tsq.12140>
- [34] Future of Privacy Forum. 2018. *Privacy Principles for Facial Recognition Technology in Commercial Applications*. <https://fpf.org/wp-content/uploads/2019/03/Final-Privacy-Principles-Edits-1.pdf>
- [35] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III au2, and Kate Crawford. 2020. Datasheets for Datasets. arXiv:1803.09010 [cs.DB]
- [36] Gender Spectrum. 2019. *The Language of Gender*. <https://genderspectrum.org/articles/language-of-gender>
- [37] Shona Ghosh. 2019. *Google AI will no longer use gender labels like ‘woman’ or ‘man’ on images of people to avoid bias*. <https://www.businessinsider.com/google-cloud-vision-api-wont-tag-images-by-gender-2020-2>
- [38] W. Gies, J. Overby, N. Saraceno, J. Frome, E. York, and A. Salman. 2020. Restricting Data Sharing and Collection of Facial Recognition Data by the Consent of the User: A Systems Analysis. In *2020 Systems and Information Engineering Design Symposium (SIEDS)*. 1–6. <https://doi.org/10.1109/SIEDS49339.2020.9106661>
- [39] Cole Gleason, Patrick Carrington, Cameron Cassidy, Meredith Ringel Morris, Kris M. Kitani, and Jeffrey P. Bigham. 2019. “It’s Almost like They’re Trying to Hide It”: How User-Provided Image Descriptions Have Failed to Make Twitter Accessible. In *The World Wide Web Conference (San Francisco, CA, USA) (WWW ’19)*. Association for Computing Machinery, New York, NY, USA, 549–559. <https://doi.org/10.1145/3308558.3313605>
- [40] Cole Gleason, Amy Pavel, Emma McCamey, Christina Low, Patrick Carrington, Kris M. Kitani, and Jeffrey P. Bigham. 2020. *Twitter A11y: A Browser Extension to Make Twitter Images Accessible*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376728>
- [41] Neil Gotanda. 1991. A Critique of “Our Constitution Is Color-Blind”. *Stanford Law Review* 44, 1 (1991), 1–68. <http://www.jstor.org/stable/1228940>
- [42] Martin Grayson, Anja Thieme, Rita Marques, Daniela Massiceti, Ed Cutrell, and Cecily Morrison. 2020. A Dynamic AI System for Extending the Capabilities of Blind People. In *CHI 2020 Extended Abstracts*. ACM. <https://www.microsoft.com/en-us/research/publication/a-dynamic-ai-system-for-extending-the-capabilities-of-blind-people/>
- [43] Darren Guinness, Edward Cutrell, and Meredith Ringel Morris. 2018. Caption Crawler: Enabling Reusable Alternative Text Descriptions Using Reverse Image Search. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI ’18*). Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3174092>
- [44] Anhong Guo, Ece Kamar, Jennifer Wortman Vaughan, Hanna Wallach, and Meredith Ringel Morris. 2020. Toward Fairness in AI for People with Disabilities: A Research Roadmap. *SIGACCESS Access. Comput.* 125, Article 2 (March 2020), 1 pages. <https://doi.org/10.1145/3386296.3386298>
- [45] Jessica Guynn. 2020. *You can now have your very own digital avatar on Facebook and Messenger*. <https://www.usatoday.com/story/tech/2020/05/13/facebook-messenger-avatars-bitmoji-snapchat-emoji/3118525001/>
- [46] Anthony Ha. 2020. *Microsoft’s Brad Smith says company will not sell facial recognition tech to police*. <https://techcrunch.com/2020/06/11/microsofts-brad-smith-says-company-will-not-sell-facial-recognition-tech-to-police/>
- [47] Stephanie Hackett, Bambang Parmanto, and Xiaoming Zeng. 2003. Accessibility of Internet Websites through Time. In *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility* (Atlanta, GA, USA) (*Assets ’04*). Association for Computing Machinery, New York, NY, USA, 32–39. <https://doi.org/10.1145/1028630.1028638>
- [48] Stephanie Hackett, Bambang Parmanto, and Xiaoming Zeng. 2005. A retrospective look at website accessibility over time. *Behaviour & Information Technology* 24, 6 (2005), 407–417. <https://doi.org/10.1080/01449290500066661>
- [49] Ever Lee Hairston, Denice Brown, Ron Brown, Bobbi Pompei, and Tarik Williams. 2020. *Not Blind to Color in the Federation: A Panel on the Experience of Black and Blind in America*. <https://www.nfb.org/images/nfb/publications/bm/bm20/bm2008/bm200813.htm>
- [50] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. 2018. Gender Recognition or Gender Reductionism? The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI ’18*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173582>
- [51] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. Towards a Critical Race Methodology in Algorithmic Fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) (*FAT* ’20*). Association for Computing Machinery, New York, NY, USA, 501–512. <https://doi.org/10.1145/3351095.3372826>
- [52] Taylor Hatmaker. 2020. *Portland passes expansive city ban on facial recognition tech*. <https://techcrunch.com/2020/09/09/facial-recognition-ban-portland-oregon/>
- [53] Anna Lauren Hoffmann. [n.d.]. Terms of inclusion: Data, discourse, violence. *New Media & Society* 0, 0 ([n. d.]), 1461444820958725. <https://doi.org/10.1177/1461444820958725>
- [54] Anna Lauren Hoffmann. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society* 22, 7 (2019), 900–915. <https://doi.org/10.1080/1369118X.2019.1573912>
- [55] Megan Hofmann, Devva Kasnitz, Jennifer Mankoff, and Cynthia L Bennett. 2020. Living Disability Theory: Reflections on Access, Research, and Design. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) (*ASSETS ’20*). Association for Computing Machinery, New York, NY, USA, Article 4, 13 pages. <https://doi.org/10.1145/3373625.3416996>

- [56] Noelle J Hum, Perrin E Chamberlin, Brittany L Hambright, Anne C Portwood, Amanda C Schat, and Jennifer L Bevan. 2011. A picture is worth a thousand words: A content analysis of Facebook profile photographs. *Computers in Human Behavior* 27, 5 (2011), 1828–1833. <https://doi.org/10.1016/j.chb.2011.04.003>
- [57] Margaret Hunter. 2007. The Persistent Problem of Colorism: Skin Tone, Status, and Inequality. *Sociology Compass* 1, 1 (2007), 237–254. <https://doi.org/10.1111/j.1751-9020.2007.00006.x>
- [58] IBM. 2020. *IBM CEO's Letter to Congress on Racial Justice Reform*. <https://www.ibm.com/blogs/policy/facial-recognition-sunset-racial-justice-reforms/>
- [59] IBM Research. 2020. *2018 Workshop on AI Fairness for People with Disabilities*. https://researcher.watson.ibm.com/researcher/view_group.php?id=9642
- [60] James M Jones. 1997. *Prejudice and racism*. McGraw-Hill Companies, New York.
- [61] Alison Kafer. 2013. *Feminist, Queer, Crip*. Indiana University Press. <http://www.jstor.org/stable/j.ctt16gz79x>
- [62] Sanja Kapidzic and Susan C Herring. 2015. Race, gender, and self-presentation in teen profile photographs. *New Media & Society* 17, 6 (2015), 958–976. <https://doi.org/10.1177/1461444813520301>
- [63] Ellyn Kaschak. 2015. *Sight Unseen*. Columbia University Press. <https://doi.org/10.7312/kasc17290>
- [64] Os Keyes. 2018. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 88 (Nov. 2018), 22 pages. <https://doi.org/10.1145/3274357>
- [65] Sooyeon Lee, Madison Reddie, Chun-Hua Tsai, Jordan Beck, Mary Beth Rosson, and John M. Carroll. 2020. *The Emerging Professional Practice of Remote Sighted Assistance for People with Visual Impairments*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376591>
- [66] Talila A. Lewis. 2020. *Ableism 2020: An Updated Definition*. <https://www.talilalewis.com/blog/ableism-2020-an-updated-definition>
- [67] Veronica Lewis. 2020. *How to write alt text and image descriptions for photojournalism images*. <https://veroniiiica.com/2020/05/31/how-to-write-alt-text-and-image-descriptions-for-photojournalism-images/>
- [68] Haley MacLeod, Cynthia L. Bennett, Meredith Ringel Morris, and Edward Cutrell. 2017. *Understanding Blind People's Experiences with Computer-Generated Captions of Social Media Images*. Association for Computing Machinery, New York, NY, USA, 5988–5999. <https://doi.org/10.1145/3025453.3025814>
- [69] Christine Malec and JJ Hunt. 2020. *Talk Description to Me*. <https://talkdescriptiontome.buzzsprout.com/>
- [70] Microsoft Research. 2017. *Seeing AI app – Scene Channel*. https://youtu.be/4krZgdX_WT8
- [71] Ryan A Miller. 2017. “My voice is definitely strongest in online communities”: Students using social media for queer and disability identity-making, 509–525 pages. <https://doi.org/10.1353/csd.2017.0040>
- [72] Meredith Ringel Morris. 2020. AI and Accessibility. *Commun. ACM* 63, 6 (May 2020), 35–37. <https://doi.org/10.1145/3356727>
- [73] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. 2018. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3173633>
- [74] Meredith Ringel Morris, Annuska Zolyomi, Catherine Yao, Sina Bahram, Jeffrey P. Bigham, and Shaun K. Kane. 2016. “With Most of It Being Pictures Now, I Rarely Use It”: Understanding Twitter’s Evolving Accessibility to Blind Users. Association for Computing Machinery, New York, NY, USA, 5506–5516. <https://doi.org/10.1145/2858036.2858116>
- [75] Paul Mozur. 2019. *One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority*. <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>
- [76] Katie Murphy. 2018. *An Audio Descriptor Responds to “On Black Panther Audio Description – Race, Selection & Time”*. <https://disarts.org/an-audio-descriptor-responds-to-on-black-panther-audio-description-race-selection-time/>
- [77] Karen Nakamura. 2019. My Algorithms Have Determined You’re Not Human: AI-ML, Reverse Turing-Tests, and the Disability Experience. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) (ASSETS '19). Association for Computing Machinery, New York, NY, USA, 1–2. <https://doi.org/10.1145/3308561.3353812>
- [78] National Center on Disability and Journalism. 2020. *Disability Language Style Guide*. <https://ncdj.org/style-guide/>
- [79] Netflix. 2020. *Black Lives Matter recommendations*. <https://www.netflix.com/browse/genre/81299227>
- [80] Safiya Umoja Noble. 2018. Algorithms of oppression: How search engines reinforce racism. , xv, 229–xv, 229 pages.
- [81] Osagie Obasogie. 2010. Do Blind People See Race? Social, Legal, and Theoretical Considerations. *Law & society review* 44 (09 2010), 585–616. <https://doi.org/10.1111/j.1540-5893.2010.00417.x>
- [82] Ihudiya Finda Ogbonnaya-Ogburu, Angela D.R. Smith, Alexandra To, and Kentaro Toyama. 2020. *Critical Race Theory for HCI*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3313831.3376392>
- [83] Katrina M Overby. 2019. Doin’It for the Culture: Defining Blackness, Culture, and Identity on Black Twitter. (2019).
- [84] Helen Petrie, Chandra Harrison, and Sundeep Dev. 2005. Describing images on the web: a survey of current practice and prospects for the future. *Proceedings of Human Computer Interaction International (HCII)* 71 (2005), 2.
- [85] John R. Porter, Kiley Sobel, Sarah E. Fox, Cynthia L. Bennett, and Julie A. Kientz. 2017. Filtered Out: Disability Disclosure Practices in Online Dating Communities. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 87 (Dec. 2017), 13 pages. <https://doi.org/10.1145/3134722>

- [86] Halley Profitta, Reem Albaghl, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. *The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use*. Association for Computing Machinery, New York, NY, USA, 4884–4895. <https://doi.org/10.1145/2858036.2858130>
- [87] Protest Access. 2020. *ProtestAccess - Twitter*. <https://twitter.com/protestaccess>
- [88] Yolanda A. Rankin and Jakita O. Thomas. 2019. Straighten up and Fly Right: Rethinking Intersectionality in HCI Research. *Interactions* 26, 6 (Oct. 2019), 64–68. <https://doi.org/10.1145/3363033>
- [89] Thomas Reid. 2018. *Reid My Mind Radio: On Black Panther Audio Description – Race, Selection & Time*. <http://reidmymind.com/reid-my-mind-radio-on-black-panther-audio-description-race-selection-time/>
- [90] Sander De Ridder and Sofie Van Bauwel. 2013. Commenting on pictures: Teens negotiating gender and sexualities on social networking sites. *Sexualities* 16, 5–6 (2013), 565–586. <https://doi.org/10.1177/1363460713487369>
- [91] Kathryn E. Ringland, Christine T. Wolf, LouAnne E. Boyd, Mark S. Baldwin, and Gillian R. Hayes. 2016. Would You Be Mine: Appropriating Minecraft as an Assistive Technology for Youth with Autism. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) (ASSETS '16). Association for Computing Machinery, New York, NY, USA, 33–41. <https://doi.org/10.1145/2982142.2982172>
- [92] Wendy D Roth. 2016. The multiple dimensions of race. *Ethnic and Racial Studies* 39, 8 (2016), 1310–1338.
- [93] Elliot Salisbury, Ece Kamar, and Meredith Morris. 2017. Toward scalable social alt text: Conversational crowdsourcing as a tool for refining vision-to-language technology for the blind. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 5.
- [94] Sigal Samuel. 2019. *Some AI just shouldn't exist*. <https://www.vox.com/future-perfect/2019/4/19/18412674/ai-bias-facial-recognition-black-gay-transgender>
- [95] Morgan Klaus Scheuerman, Jacob M. Paul, and Jed R. Brubaker. 2019. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 144 (Nov. 2019), 33 pages. <https://doi.org/10.1145/3359246>
- [96] Morgan Klaus Scheuerman, Katta Spiel, Oliver L. Haimson, Foad Hamidi, and Stacy M. Branham. 2020. *HCI Guidelines for Gender Equity and Inclusivity*. <https://www.morgan-klaus.com/gender-guidelines.html>
- [97] Morgan Klaus Scheuerman, Kandrea Wade, Caitlin Lustig, and Jed R. Brubaker. 2020. How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 058 (May 2020), 35 pages. <https://doi.org/10.1145/3392866>
- [98] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. *Intersectional HCI: Engaging Identity through Gender, Race, and Class*. Association for Computing Machinery, New York, NY, USA, 5412–5427. <https://doi.org/10.1145/3025453.3025766>
- [99] Ari Schlesinger, Kenton P. O'Hara, and Alex S. Taylor. 2018. Let's Talk About Race: Identity, Chatbots, and AI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3173574.3173889>
- [100] Cristina Segalin, Fabio Celli, Luca Polonio, Michal Kosinski, David Stillwell, Nicu Sebe, Marco Cristani, and Bruno Lepri. 2017. What Your Facebook Profile Picture Reveals about Your Personality. In *Proceedings of the 25th ACM International Conference on Multimedia* (Mountain View, California, USA) (MM '17). Association for Computing Machinery, New York, NY, USA, 460–468. <https://doi.org/10.1145/3123266.3123331>
- [101] Sins Invalid. 2020. *An Unshamed Claim to Beauty in the Face of Invisibility*. <https://www.sinsinvalid.org/>
- [102] Olivia Solon. 2020. *Facial recognition's 'dirty little secret': Millions of online photos scraped without consent*. <https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921>
- [103] Katta Spiel, Os Keyes, Ashley Marie Walker, Michael A. DeVito, Jeremy Birnholtz, Emeline Brûlé, Ann Light, Pinar Barlas, Jean Hardy, Alex Ahmed, Jennifer A. Rode, Jed R. Brubaker, and Gopinaath Kannabiran. 2019. Queer(Ing) HCI: Moving Forward in Theory and Practice. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI EA '19). Association for Computing Machinery, New York, NY, USA, 1–4. <https://doi.org/10.1145/3290607.3311750>
- [104] Abigale Stangl, Meredith Ringel Morris, and Danna Gurari. 2020. “Person, Shoes, Tree. Is the Person Naked?” What People with Vision Impairments Want in Image Descriptions. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376404>
- [105] Abigale J. Stangl, Esha Kothari, Suyog D. Jain, Tom Yeh, Kristen Grauman, and Danna Gurari. 2018. BrowseWithMe: An Online Clothes Shopping Assistant for People with Visual Impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) (ASSETS '18). Association for Computing Machinery, New York, NY, USA, 107–118. <https://doi.org/10.1145/3234695.3236337>
- [106] Lee Stearns and Anja Thieme. 2018. Automated Person Detection in Dynamic Scenes to Assist People with Vision Impairments: An Initial Investigation. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) (ASSETS '18). Association for Computing Machinery, New York, NY, USA, 391–394. <https://doi.org/10.1145/3234695.3241017>
- [107] Anselm Strauss and Juliet Corbin. 1998. *Basics of qualitative research techniques*. Sage publications Thousand Oaks, CA.
- [108] Derald Wing Sue. 2010. *Microaggressions in everyday life: Race, gender, and sexual orientation*. John Wiley & Sons.
- [109] The Amazon blog. 2020. *We are implementing a one-year moratorium on police use of Rekognition*. <https://blog.aboutamazon.com/policy/we-are-implementing-a-one-year-moratorium-on-police-use-of-rekognition>
- [110] Matt Thompson. 2013. *Five Reasons Why People Code-Switch*. <https://www.npr.org/sections/codeswitch/2013/04/13/177126294/five-reasons-why-people-code-switch>
- [111] Jutta Treviranus. 2020. *We Count: Fair Treatment, Disability and Machine Learning - by Jutta Treviranus (OCAD University)*. https://www.w3.org/2020/06/machine-learning-workshop/talks/we_count_fair_treatment_disability_and_machine_learning.html

- [112] Shari Trewin. 2018. AI fairness for people with disabilities: Point of view. *arXiv preprint arXiv:1811.10670* (2018).
- [113] Shari Trewin, Meredith Ringel Morris, Shiri Azenkot, Stacy Branham, Nicole Bleuel, Phill Jenkins, Jeffrey P. Bigham, and Walter S. Lasecki. 2019. *Workshop: AI Fairness for People with Disabilities*. https://assets19.sigaccess.org/ai_fairness_workshop.html
- [114] James Vincent. 2018. *Amazon reportedly scraps internal AI recruiting tool that was biased against women*. <https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report>
- [115] Violeta Voykinska, Shiri Azenkot, Shaomei Wu, and Gilly Leshed. 2016. How Blind People Interact with Visual Content on Social Networking Services. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (*CSCW '16*). Association for Computing Machinery, New York, NY, USA, 1584–1595. <https://doi.org/10.1145/2818048.2820013>
- [116] W3C. 2018. *Web Content Accessibility Guidelines (WCAG) 2.1*. <https://www.w3.org/TR/WCAG21/>
- [117] W3C. 2019. *Informative Images • Images • WAI Web Accessibility Tutorials*. <https://www.w3.org/WAI/tutorials/images/informative/>
- [118] Sarah Myers West, Meredith Whittaker, and Kate Crawford. 2019. *DISCRIMINATING SYSTEMS: Gender, Race, and Power in AI*. <https://ainowinstitute.org/discriminatingsystems.html>
- [119] Meredith Whittaker, Meryl Alper, Cynthia L Bennett, Sara Hendren, Liz Kaziunas, Mara Mills, Meredith Ringel Morris, Joy Rankin, Emily Rogers, Marcel Salas, et al. 2019. Disability, Bias, and AI. *AI Now Institute, November* (2019).
- [120] Rua M. Williams and Juan E. Gilbert. 2019. Cyborg Perspectives on Computing Research Reform. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (*CHI EA '19*). Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3290607.3310421>
- [121] Writing with Color. 2021. *Writing with Color*. <https://writingwithcolor.tumblr.com/>
- [122] Shaomei Wu and Lada A. Adamic. 2014. Visually Impaired Users on an Online Social Network. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 3133–3142. <https://doi.org/10.1145/2556288.2557415>
- [123] Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic Alt-Text: Computer-Generated Image Descriptions for Blind Users on a Social Network Service. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (*CSCW '17*). Association for Computing Machinery, New York, NY, USA, 1180–1192. <https://doi.org/10.1145/2998181.2998364>
- [124] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. 2017. The Effect of Computer-Generated Descriptions on Photo-Sharing Experiences of People with Visual Impairments. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 121 (Dec. 2017), 22 pages. <https://doi.org/10.1145/3134756>
- [125] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. 2018. A Face Recognition Application for People with Visual Impairments: Understanding Use Beyond the Lab. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3173574.3173789>
- [126] Wenzhi Zheng, Chih-Hung Yuan, Wei-Hung Chang, and Yen-Chun Jim Wu. 2016. Profile pictures on social media: Gender and regional differences. *Computers in Human Behavior* 63 (2016), 891–898.